Two Perspectives on ML

Computer science vs. Statistical perspective

As you can see, data plays a central role in how problems are modeled in machine learning. In very broad terms, we can think of machine learning as a matter of using some data (perhaps historical data that we already have on hand) to train a model. Then, once the model is trained, we can feed it new input data and have it tell us something useful.

So the general idea is that we create models and then feed data into these models to generate outputs. These outputs might be, for example, predictions for future trends or patterns in the data.

This idea draws on work not only from *computer science*, but also *statistics*—and as a result, you will often see the same underlying machine learning concepts described using different terms. For example, a computer scientist might say something like:

> *We are using **input features** to create a **program** that can generate the desired **output**.*

In contrast, someone with a background in statistics might be inclined to say something more like:

> *We are trying to find a **mathematical function** that, given the values of the **independent variables** can predict the values of the **dependent variables**.*

While the terminology are different, the challenges are the same, that is how to get the best possible outcome.

QUIZ QUESTION

Can you match the terms below, from the computer science perspective, with their counterparts from the statistical perspective?

Submit to check your answer choices!

COMPUTER SCIENCE	STATISTICAL
program	function
input	independent variable
output	dependent variable

SUBMIT

In the end, having an understanding of the underlying concepts is more important than memorizing the terms used to describe those concepts. However, it's still essential to be familiar with the terminology so that you don't get confused when talking with people from different backgrounds.

Over the next couple of pages, we'll take a look at these two different perspectives and get familiar with some of the related terminology.